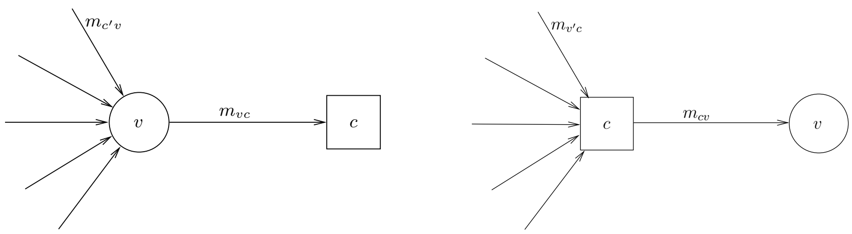
Lecture 19

Properties of Belief Propagation I

In this lecture we consider transmission over a binary input, memoryless symmetric channel (BIMS), using an LDPC code chosen randomly from the ensemble $\mathcal{C}(\lambda, \rho, n)$ of codes of blocklength n with left and right edge degree distribution λ and ρ, respectively. Throughout we will use binary antipodal signaling, i.e., $0 \mapsto 1$ and $1 \mapsto -1$.

19.1. Belief-Propagation

At each iteration of the decoding algorithm messages are passed from variable nodes to incident check nodes and back. Let us denote by m_{vc} the message passed from variable node v to check node c, and let m_{cv} be the message passed from check node c to variable node v (see Figure 19.1). Each message m represents the log-likelihood ratio (LLR) of the codeword bit whose associated variable note sends or receives the message, given all information that was used in computing the message.



(a) Message from variable node to check node

(b) Message from check node to variable node

Figure 19.1: Edge messages in the belief propagation algorithm

At the first step of the algorithm, each variable node v sends m_v, the LLR of the corresponding codeword bit x given only the channel output observation y:

$$m_v = \ln \frac{\Pr[x = +1|y]}{\Pr[x = -1|y]}$$

Lecture by Amin Shokrollahi
Scribes by Marius Kleiner

To represent the updates performed at check nodes in a concise form, it will be useful to introduce a map γ, defined as

$$\begin{aligned} \gamma : \mathbb{R} &\rightarrow \mathbb{F}_2 \times \mathbb{R}_{\geq 0} \\ x &\mapsto \gamma(x) = (\gamma_1(x), \gamma_2(x)) = (\operatorname{sgn}(x), -\ln\tanh(|x|/2)), \end{aligned} \tag{19.1}$$

where

$$\operatorname{sgn}(x) = \begin{cases} 0 & \text{if } x \geq 0 \\ 1 & \text{if } x < 0 \end{cases} .$$

Note that γ is a bijective map, its inverse is $\gamma^{-1}(s, y) = 2(-1)^s \tanh^{-1}(e^{-y})$.

Recall from Lecture 13 the update rules for the belief-propagation algorithm at variable nodes and check nodes:

$$m_{vc} = \sum_{c' \in C_v \setminus \{c\}} m_{c'v} + m_v \tag{19.2}$$

$$\tanh\left(\frac{m_{cv}}{2}\right) = \prod_{v' \in V_c \setminus \{v\}} \tanh\left(\frac{m_{v'c}}{2}\right), \tag{19.3}$$

where C_v is the set of check nodes incident to variable node v, and V_c is the set of variable nodes incident to check node c. Using the map γ we can write the update rule for check nodes (19.3) as

$$m_{cv} = \gamma^{-1}\left(\sum_{v' \in V_c \setminus \{v\}} \gamma(m_{v'c})\right). \tag{19.4}$$

19.2. Density Evolution

We will now analyze how the probability distributions of individual messages evolve during successive iterations of the belief-propagation algorithm. Suppose that at iteration i we randomly pick an edge from the code graph. Let m_{vc}^i be the message that is sent across the edge from the variable node to the check node, and let f_i be its density, i.e, $\int_{-\infty}^{x} f_i(z)\mathrm{d}z = \Pr[m_{vc}^i \leq x]$. The expression for m_{vc}^i is given by (19.2), and assuming the incoming messages are statistically independent, we have

$$f_i = f_C \otimes g_{i-1}^{\otimes(d-1)},$$

where f_C is the density of the observation LLR at the variable node, g_{i-1} are the densities of the incoming messages from the check nodes at the previous iteration, and d is the degree of the variable node. Since we chose an edge at random, we need to average over all possible degrees d, obtaining

$$f_i = f_C \otimes \sum_k \lambda_k g_{i-1}^{\otimes(k-1)}.$$

With a slight but convenient abuse of notation we define the "convolution polynomial" $\lambda(x) = \sum_k \lambda_k x^{\otimes(k-1)}$ and write

$$f_i = f_C \otimes \lambda(g_{i-1}). \tag{19.5}$$

As for the message from the check node to the variable node, we cannot directly find an expression departing from equation (19.3) since there is no general formula for the density of a product of random variables. If we could express the densities of the incoming messages under the transformation γ, however, we could apply the convolution just as for the variable node.

To compute these densities, consider a random variable $X \in \mathbb{R}$ whose cumulative distribution function (CDF) is F, i.e., $F(x) = \Pr[X \leq x]$. We want to compute the distribution of the transformed random variable $\gamma(X) \in \mathbb{F}_2 \times \mathbb{R}_{\geq 0}$. Let us denote this distribution by $\Gamma(F)(s, y)$. Since its first argument takes just two possible values, 0 or 1, it can be written as

$$\Gamma(F)(s, y) = \chi_{\{s=0\}} G_0(y) + \chi_{\{s=1\}} G_1(y), \tag{19.6}$$

where $\chi_{\{s=a\}}$ is the characteristic function of the event $\{s = a\}$. When $s = 0$, i.e., $X \geq 0$, we have

$$\begin{aligned} G_0(y) &= \Pr[-\ln\tanh(X/2) \leq y] \\ &= \Pr\left[\tanh(X/2) \geq e^{-y}\right] \\ &= \Pr\left[\frac{e^X - 1}{e^X + 1} \geq e^{-y}\right] \\ &= \Pr\left[e^X \geq \frac{1 + e^{-y}}{1 - e^{-y}}\right] \\ &= \Pr\left[X \geq -\ln\tanh(y/2)\right] \\ &= 1 - F(-\ln\tanh(y/2)). \end{aligned} \tag{19.7}$$

Similarly, when $s = 1$ we have

$$\begin{aligned} G_1(y) &= \Pr[-\ln\tanh(-X/2) \leq y] \\ &= \Pr[\tanh(X/2) \leq -e^{-y}] \\ &= \Pr\left[\frac{e^X - 1}{e^x + 1} \leq -e^{-y}\right] \\ &= \Pr\left[e^X \leq \frac{1 - e^{-y}}{1 + e^{-y}}\right] \\ &= \Pr[X \leq \ln\tanh(y/2)] \\ &= F(\ln\tanh(y/2)). \end{aligned} \tag{19.8}$$

Note at this point that Γ is invertible. Indeed, if $G = \chi_{\{s=0\}}G_0 + \chi_{\{s=1\}}G_1$ is a distribution on $\mathbb{F}_2 \times \mathbb{R}_{\geq 0}$, we have

$$\Gamma^{-1}(G)(x) = \chi_{\{x\geq 0\}}G_0\left(-\ln\tanh(x/2)\right) + \chi_{\{x<0\}}G_1\left(-\ln\tanh(-x/2)\right).$$

Taking derivatives of (19.7) and (19.8) to obtain densities, we get

$$\frac{\mathrm{d}}{\mathrm{d}y}G_0(y) = \frac{f(-\ln\tanh(y/2))}{\sinh(y)} \tag{19.9}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}y}G_1(y) = \frac{f(\ln\tanh(y/2))}{\sinh(y)}. \tag{19.10}$$

Let us now turn back to density evolution. Recall the update rule for check nodes (19.4):

$$m_{cv} = \gamma^{-1}\left(\sum_{v'\in V_c\setminus\{v\}} \gamma(m_{v'c})\right).$$

We shall denote by f_i the densities of the incoming messages $m_{v'c}$. We allow ourselves to write $\Gamma(f_i)$ for the densities of the transformed messages $\gamma(m_{v'c})$, reminding the reader that behind the scenes Γ is actually applied to the *distribution* as in (19.6). We can now apply the convolution property to the densities $\Gamma(f_i)$ of the (transformed) incoming messages. In fact, if X and Y are two random variables on $\mathbb{F}_2 \times \mathbb{R}_{\geq 0}$ with respective densities $f = \chi_{\{s=0\}}f_0 + \chi_{\{s=1\}}f_1$ and $g = \chi_{\{s=0\}}g_0 + \chi_{\{s=1\}}g_1$, the density of $X+Y$ is $\chi_{\{s=0\}}((f_0 \otimes g_0) + (f_1 \otimes g_1)) + \chi_{\{s=1\}}((f_0 \otimes g_1) + (f_1 \otimes g_0))$, where here $\otimes$ denotes the (one–sided) convolution of standard densities. Let us again use $\otimes$ to denote this new convolution over $\mathbb{F}_2 \times \mathbb{R}_{\geq 0}$. Thus, for a check node of degree d, we have from (19.4)

$$g_i = \Gamma^{-1}\left(\Gamma(f_i)^{\otimes(d-1)}\right).$$

Taking the expectation over the degree distribution we get

$$g_i = \Gamma^{-1}\left(\sum_k \rho_k \Gamma(f_i)^{\otimes(k-1)}\right),$$

and as before we define a "convolution polynomial" $\rho(x) = \sum_k \rho_k x^{\otimes(k-1)}$ to obtain

$$g_i = \Gamma^{-1}\left(\rho\left(\Gamma(f_i)\right)\right). \tag{19.11}$$

Combining (19.11) with (19.5) we obtain the final density evolution formula

$$f_i = f_C \otimes \lambda\left(\Gamma^{-1}\left(\rho\left(\Gamma(f_{i-1})\right)\right)\right). \tag{19.12}$$

19.3. Application to the Binary Erasure Channel

The density evolution formula (19.12) is valid for any BIMS channel, in particular for the binary erasure channel (BEC). Let us now evaluate (19.12) for the specific case of the BEC and compare the result to that obtained in Lecture 17.

Under the assumption that the all-one codeword was sent, the LLR messages passed around during decoding take only the values 0 and ∞; their distribution at iteration i consists therefore of just two mass points and can be characterized by a single value x_i, defined as

$$x_i = \Pr[m_{vc}^i = 0],$$

from which follows

$$1 - x_i = \Pr[m_{vc}^i = \infty].$$

The corresponding CDF (over the set $\mathbb{R} \cup \{\infty\}$) takes on the form

$$F_i(x) = x_i \chi_{\{x \geq 0\}}$$

and $F_i(\infty) = 1$. Since the bijective transformation γ is such that $\gamma(0) = (0, \infty)$ and $\gamma(\infty) = (0, 0)$ (cf. (19.1)), the distribution $\Gamma(F_i)$ can be written as

$$\Gamma(F_i)(s, y) = (1 - x_i)\chi_{\{s=0\}}. \tag{19.13}$$

Similarly, since γ is bijective the inverse transformation of a distribution on $\mathbb{F}_2 \times \mathbb{R}_{\geq 0}$ of the form $x_i \chi_{\{s=0\}}$ is

$$\Gamma^{-1}(G_i)(x) = (1 - x_i)\chi_{\{x \geq 0\}}. \tag{19.14}$$

Let us now recall the density evolution formula (19.12):

$$f_i = f_C \otimes \lambda\left(\Gamma^{-1}\left(\rho\left(\Gamma(f_{i-1})\right)\right)\right).$$

From (19.13) above we have

$$\Gamma(F_i)(s, y) = \chi_{\{s=0\}}(1 - x_i).$$

We then obtain

$$\rho\left(\Gamma(F_i)\right)(s, y) = \chi_{\{s=0\}}\rho(1 - x_i)$$

and from (19.14)

$$\Gamma^{-1}\left(\rho\left(\Gamma\left(F_i\right)\right)\right)(x) = \left(1 - \rho(1 - x_i)\right)\chi_{\{x \geq 0\}}.$$

Finally,

$$F_C \otimes \lambda\left(\Gamma^{-1}\left(\rho\left(\Gamma(F_i)\right)\right)\right) = x_0 \lambda\left(1 - \rho(1 - x_i)\right)\chi_{\{x \geq 0\}},$$

where x_0 is the probability that the observation LLR is zero. From this we can recover the density evolution formula

$$x_i = x_0 \lambda(1 - \rho(1 - x_{i-1})),$$

which is the same as the one found in Lecture 17.